

# Einführung in Prompt-Engineering: Die Kunst mit den KI-Modellen zu kommunizieren

Ludwig-Maximilians-Universität, München

Centrum für Informations- und Sprachverarbeitung

Masterseminar: Conversational AI

Dozent: Dr. Christoph Ringlstetter

Referentin: Evanthia Tsigkana

München, 21.11.2024



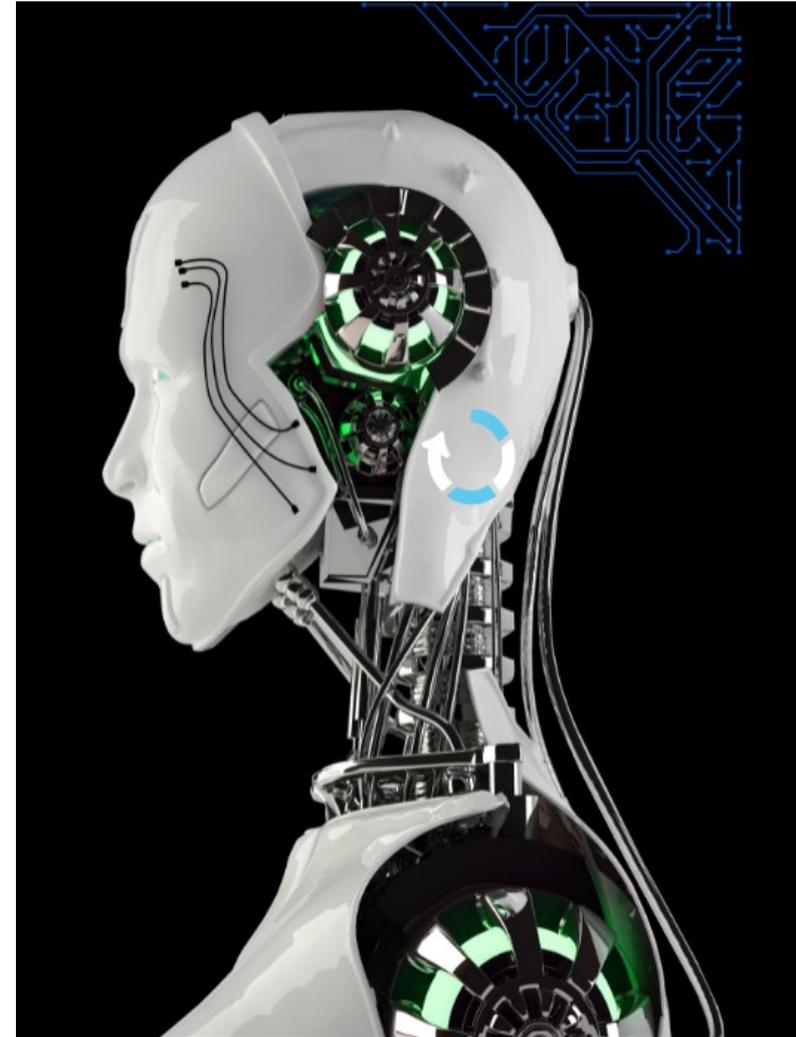
# Übersicht

1. Einleitung
2. Effektivität, Anwendung und Methoden
3. Herausforderungen und Grenzen
4. Zusammenfassung und Ausblick
5. Literaturhinweise

# Einleitung: Was ist Prompting ?

## Definition

- Fähigkeit klare und strukturierte Prompts zu formulieren.
- Erzielen von möglichst den Besten Antworten von KI-Modellen.
- Bekannt auch als In-Context Prompting.



(Abb.1 <https://medium.com/@ensargnsdogdu/unveiling-the-power-of-ai-prompt-engineering-demystifying-and-harnessing-its-potential-5cc176147951>)

# Einleitung

- Ein präzises Prompt erhöht die Genauigkeit der KI-Antwort.
- Zeit sparen durch relevante Antworten.
- Beispiele: „Social Media.“ vs. „Schreibe kurz über die Vorteile und Nachteile von Social Media und wie sie unser tägliches Leben beeinflussen.“

# Effektivität, Anwendung, Methoden

Wie steigt die Effektivität von Prompting?

- Präzise Fragen stellen und relevanten Kontext hinzufügen
- Beispiele verwenden
- Komplexe Anfragen in mehreren Teilen stellen
- Das neueste Modell verwenden

# Effektivität, Anwendung, Methoden

Wo wird das Prompt-Engineering eingesetzt?

- *Kundenberatung und Chatbots* durch Automatisierung der Antworten.
- *Wissenschaftliche Forschung* zB. in der Medizin können die Modelle die wichtigsten Informationen aus der Literatur extrahieren und komplizierte Sachverhalte für alle Leser darstellen
- *Bildung und E-Learning* zB. Lernmaterialien erstellen

# Effektivität, Anwendung, Methoden

- *Software-Entwicklung.* Entwickler nutzen das Prompt-Engineering als Erklärung von Codes oder als Inspiration beim Coding.
- *Content erstellen und Marketing.* Durch spezifische Prompts werden die Modelle auf bestimmten Schreibstil und Zielgruppen abgestimmt.
- *Kreativität fördern.* Prompts werden eingesetzt zur Generierung von fiktiven Figuren und Bildern.

# Effektivität, Anwendung, Methoden

- *Datenanalyse* in sämtlichen Bereichen. Reports erstellen, Quartalszahlen analysieren, Risiken erkennen etc.

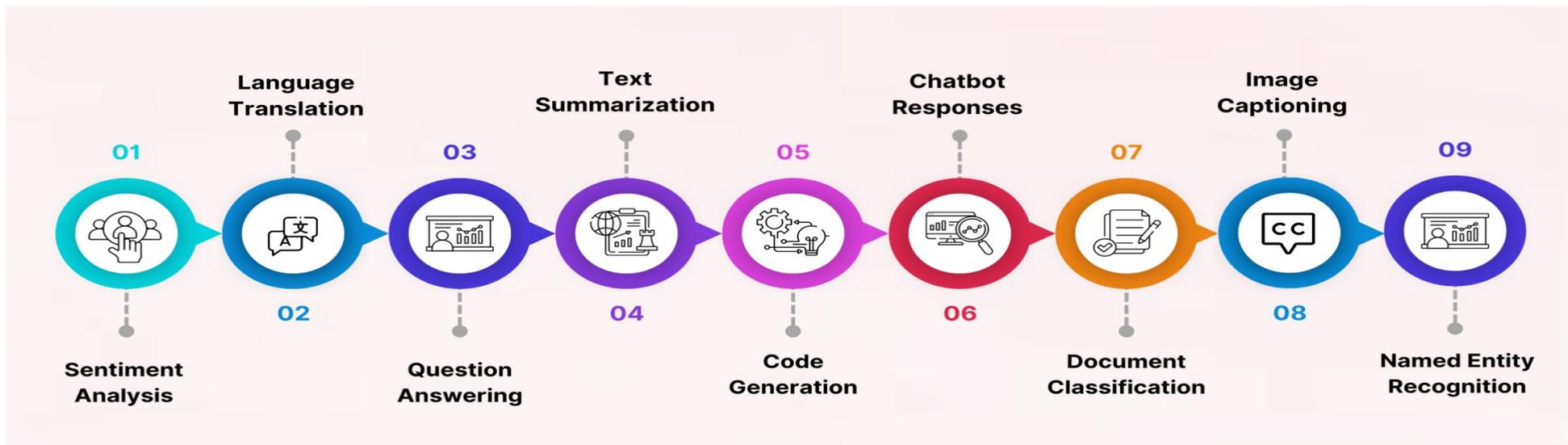


Abb.2 <https://www.a3logics.com/blog/prompt-engineering-use-cases/>

# Effektivität, Anwendung, Methoden

## Methoden:

- Beginnen mit Anweisungen und Rauten # # # oder Anführungszeichen “ “ “ zwischen Anweisung und Kontext verwenden.

Bsp.: Weniger effektiv: ***Summarize the text below as a bullet point list of the most important points.***

***(text input here)***

Effektiver: ***Summarize the text below as a bullet point of the most important points.***

***Text: ““““***

***(text input here) ““““***

# Effektivität, Anwendung, Methoden

Methoden:

- Möglichst konkrete Anweisungen geben über Kontext, Länge, Format, Stil etc.

Beispiel:

Weniger effektiv: ***Schreib mir einen Bericht über Reisen.***

Effektiver: ***Schreib mir einen kurzen Reisebericht über Italien und fokussiere auf die schönsten Strände im Süden des Landes.***

# Effektivität, Anwendung, Methoden

Methoden:

- ***Zero-Shot-Prompting*** und ***Few-Shot-Prompting***: wichtige Elemente des Prompt-Engineering in der KI

## Zero-Shot-Prompting

Eine konkrete Frage (Prompt) stellen, um eine Aufgabe zu lösen. Kein Training bzgl. der Aufgabe.

- Hilfreich bei schnellen ersten Antworten/Lösungen
- Kein umfangreiches fine-tuning des Sprachmodells

# Effektivität, Anwendung, Methoden

Methoden:

Beispiel für Zero-Shot-Prompting:

***„Beschreibe das Zero-Shot-Prompting“***

Das Modell gibt eine Antwort aus seinem allgemeinen Wissen, ohne zuvor mit ähnlichen Beispielen trainiert zu sein.

# Effektivität, Anwendung, Methoden

Methoden:

- Few-Shot-Prompting

Dem Modell ein paar Beispiel-Fragen und -  
Antworten geben, bevor die tatsächliche Frage  
gestellt wird.

→ Das Modell wird dadurch gezielt trainiert

→ Es bekommt eine bessere Vorstellung davon, was  
gefragt wird und wie die Antwort aussehen soll

# Effektivität, Anwendung, Methoden

- Das Modell nutzt die Beispiele als Lernhilfen
- Es passt seine Antwort entsprechend an
- Hilfreich bei komplexen Aufgaben
- Trainiert für zahlreiche Aufgaben, z.B. Sentiment Analysis, Übersetzung

# Effektivität, Anwendung, Methoden

## Beispiele für Few-Shot-Prompting

Wir wollen dem Model beibringen, die negative oder positive Stimmung zu erkennen

Prompt:

- Bsp. 1 „Das Wetter war schlecht.“ → Stimmung: negativ
- Bsp. 2 „Der Film war fantastisch.“ → Stimmung:  
positiv
- Tatsächliche Frage: „Das Essen war lecker.“ → Stimmung:  
Erwartete Antwort: positiv

# Effektivität, Anwendung, Methoden

- Starten mit Zero-Shot-Prompting und weitermachen mit Few-Shot.
- Manche Modelle wie ChatGPT zeigen oft mit Zero-Shot-Prompting eine schlechte Performanz.
- Bei schlechten Ergebnissen → Fine-Tuning des Sprachmodells.

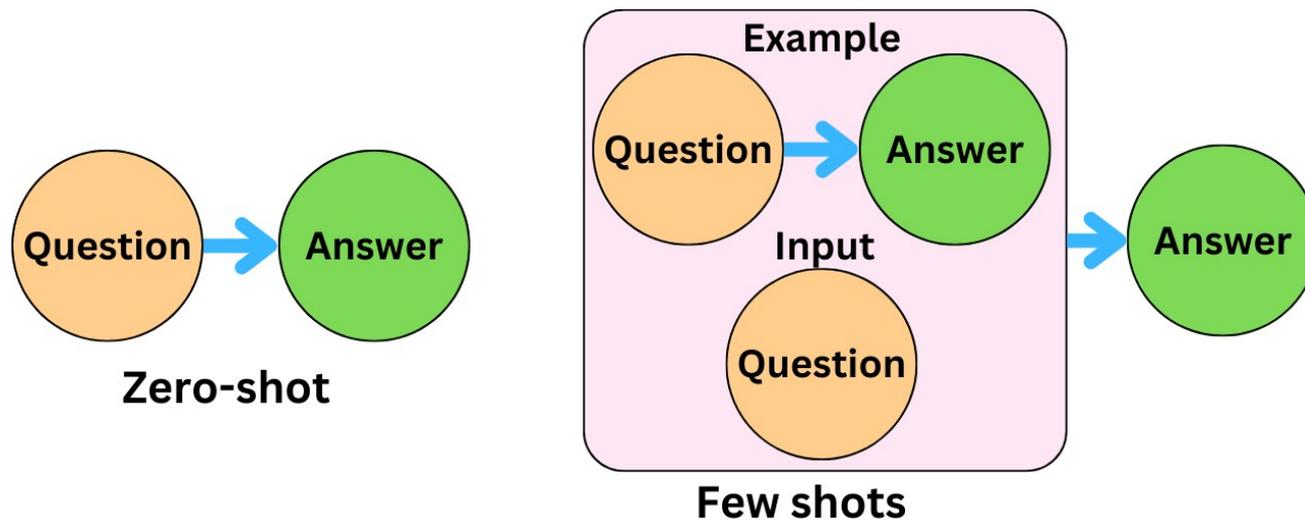


Abb. 3 <https://newsletter.theaiedge.io/p/prompt-engineering-and-llmops-building>

# Effektivität, Anwendung, Methoden

## Methoden: Chain of Thought (CoT)

- Bei Aufgaben, die logisches Denken und mehrere Schlussfolgerungen benötigen.
- Diese Technik hilft dem Modell komplexere Probleme Schritt für Schritt zu lösen.
- Eingesetzt bei mathematischen und logischen Problemen, größeren Entscheidungsprozessen, Analysen von komplexen Themen.
- Vorteilhaft wegen Genauigkeit und Präzision. Kombinierbar mit den anderen Techniken.

# Effektivität, Anwendung, Methoden

## Beispiel für **Chain of Thought (CoT)**

Problemstellung → „Berechne die Gehälter von 3 Mitarbeitern für 2 Monate.“

1. Zuerst das Gehalt von 1 Mitarbeiter für 1 Monat berechnen.
2. Jeder Mitarbeiter verdient 4.000 € netto pro Monat.
3. Jeder Mitarbeiter verdient 8.000 € netto in zwei Monaten.
4. Drei Mitarbeiter verdienen 24.000 € netto in zwei Monaten

# Effektivität, Anwendung, Methoden

## Beispiel für **Chain of Thought (CoT)**

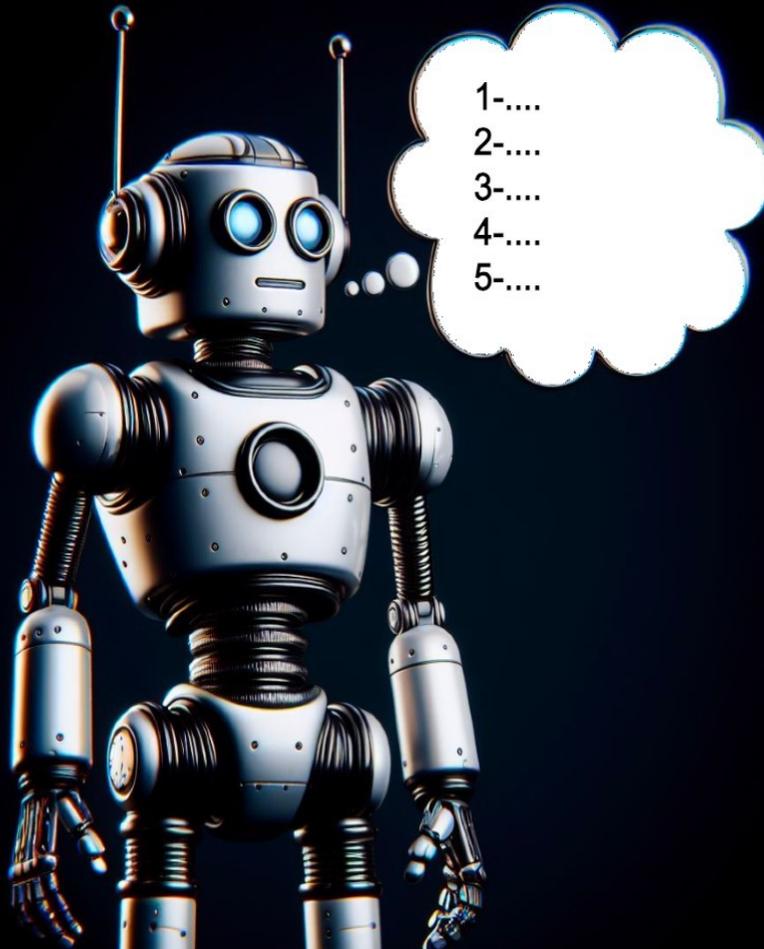


Abb. 4 <https://bdtechtalks.com/2024/05/13/chain-of-thought-planning/>

# Effektivität, Anwendung, Methoden

## Methoden: Negative Prompting

- Verwendet meistens bei Text-To-Bild-Generatoren wie DALLE oder Stable Diffusion. Auch bei NLP und Machine Learning.
- Unerwünschte Elemente oder Bias werden aus den Ergebnissen entfernt.
- **Funktionsweise:** Neben einem Prompt für etwas Erwünschtes, ein Prompt eingeben, der beschreibt, was NICHT als Ergebnis herauskommen soll. Das Model versucht die negativen Aspekte zu vermeiden.
- Methode für präzisere Ergebnisse und zum Vermeiden von Fehlern.

# Effektivität, Anwendung, Methoden

## Methoden: Negative Prompting

Beispiel für Text-to-Bild:

Wir wollen ein Bild von einer Landschaft mit Wald und bunten Farben beim Sonnenaufgang, aber keine Autos, kein Müll und keine Menschen.

**Prompt:** Eine Landschaft mit Wald und bunten Farben beim Sonnenaufgang.

**Negativer Prompt:** keine Autos, kein Müll, keine Menschen

# Effektivität, Anwendung, Methoden

## Methoden: Iterative Prompting

- Erhalten präzisere und relevantere Antworten.
- Der Benutzer stellt eine Frage oder Anweisung.
- Er verbessert die Antwort schrittweise durch Nachfragen.
- **Funktionsweise:** Starten mit einem Prompt, Prüfen die Antwort, Verbessern des Prompts, Wiederholen den Prozess.
- Vorteile: Schrittweise Präzisierung und Flexibilität.

# Effektivität, Anwendung, Methoden

## Methoden: Iterative Prompting

Beispiel: Planung einer Reise

- Anfangsprompt: Welche sind die besten Reiseziele in Europa?
- Weitere Prompts je nach Antwort und Nachhaken, bis die Antwort zufriedenstellend beantwortet ist.

# Herausforderungen und Grenzen

- Ethische Herausforderungen. Stereotypische Antworten bei sozialen und gesellschaftlichen Themen. Auf unkontrollierte Datensätze trainierte Modelle.
- Mangelndes aktuelles Fachwissen. Die Modelle sind nicht immer auf die aktuellen Themen trainiert.
- Ambiguität und Missverständnisse. Mehrdeutige Prompts führen dazu, dass sich das Modell für eine Interpretation entscheidet. Eventuell falsch.

# Herausforderungen und Grenzen

- Komplexe Aufgaben. Prompts zu wissenschaftlichen Fragen werden oberflächlich beantwortet.
- Menschliche Emotionen und Nuancen werden missverstanden.
- Zeitaufwand für die richtigen und besten Prompts.
- Prompt Overfitting → Verlust der Generalisierung, Zeitaufwand für das spezifische Prompt, keine Flexibilität bei den Antworten, Einschränkung der Effizienz.

# Zusammenfassung und Ausblick

- Prompt Engineering ist die Technik, KI-Modelle durch Prompts zu steuern.
- Die Prompts sollen klar und präzise formuliert sein.
- Der Kontext und die Details bei der Formulierung helfen, eine qualitativ hochwertige Antwort zu generieren.
- Iterativer Prozess: Mehrere Schritte bis zum erwünschten Ergebnis.

# Zusammenfassung und Ausblick

- Wichtige Techniken: Zero-Shot-Prompting, Few-Shot-Prompting, Chain of Thoughts, Iterative Prompting, Negative Prompting.
- Es gibt mehrere Bereiche, wo Prompt-Engineering eingesetzt wird.
- Hilfreich aber auch „gefährlich“, weil die verwendeten Daten unkontrolliert sind.
- Die Entwicklung von KI wird neue Techniken und Möglichkeiten für das Prompt Engineering bieten.

# Literaturhinweise

- A Comprehensive Overview of Large Language Models. Humza Naveed, Asad Ullah Khan, Shi Qiu, Muhammad Saqib, Saeed Anwar, Muhammad Usman, Naveed Akhtar, Nick Barnes, Ajmal Mian. arXiv:2307.06435 [cs.CL], 17 Oct2024.
- A Survey of Large Language Models, Wayne Xin Zhao et al.
- Finetuned Language Models Are Zero-Shot Learners. Jason Wei, Maarten Bosma, Vincent Y. Zhao, Kelvin Guu, Adams Wei Yu, Brian Lester, Nan Du, Andrew M. Dai, Quoc V. Le. arXiv\_2109.01652v5 [cs.CL], 8 Feb 2022.
- <https://medium.com/@amiraryani/8-types-of-prompt-engineering-5322fff77bdf>
- <https://bdtechtalks.com/2024/05/13/chain-of-thought-planning/>
- <https://www.a3logics.com/blog/prompt-engineering-use-cases/>
- <https://openai.com/index/learning-to-reason-with-llms/>
- <https://github.com/lilianweng/lilianweng.github.io/blob/master/posts/2023-03-15-prompt-engineering/index.html>
- <https://help.openai.com/en/articles/6654000-best-practices-for-prompt-engineering-with-the-openai-api>

# Vielen Dank für Ihre Aufmerksamkeit!

